Relations Between the Voice and the Ear With Clinical Implications

Eric J. Hunter
National Center for Voice and Speech,
Division of Otolaryngology-Head and Neck Surgery, Department of Communication Sciences and Disorders and Department of Bioengineering
The University of Utah
Salt Lake City, UT

Brian B. Monson
National Center for Voice and Speech
Salt Lake City, UT
and
Department of Speech, Language, and Hearing Science, University of Arizona
Tucson, AZ

Doug W. Montequin
National Center for Voice and Speech
Denver Public Schools
Denver, CO

Abstract

The ear and the vocal mechanism are closely tied. We develop our voice and speech skills with constant aural feedback by making on-line changes in our acoustic output based on our self-perceived acoustic output. The efficacy of the human voice in communicating messages and ideas to others is inevitably dependent upon the human auditory system (of both the listener and the talker). Thus, a basic understanding of the relation of these two mechanisms is fundamental to an accurate assessment of potential vocal issues. The authors will review and discuss some of the relationships between the voice and the ear, the environmental factors that can affect aural feedback in a clinical setting, and the significance of this discussion to the speech-language pathologist.

Introduction

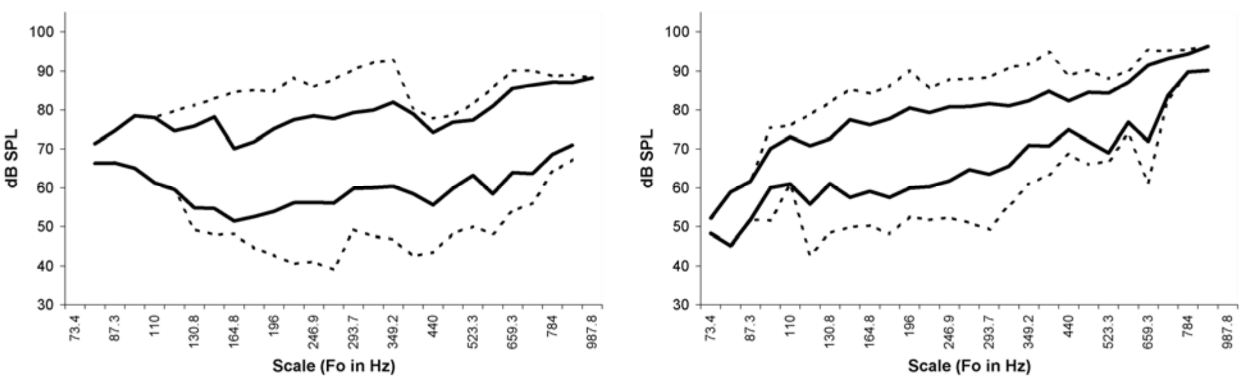
Human vocal communication relies on two mechanisms: (a) the production of voice by the vocalist and (b) the perception (or hearing) of voice by the listener. Traditionally, most clinicians treat voice and speech disorders or hearing disorders in isolation. With only a few exceptions, this same segmentation typically is seen in research as well.

However, the efficacy of the human voice in communicating ideas to others is inevitably dependent upon the human auditory system (of both the listener and the speaker). Thus, a basic understanding of the relationship between these two mechanisms is fundamental to assess the presence of a vocal issue accurately. To this end, we discuss some of the relationships between the voice and the ear. First, we review the spectral overlap of the voice and human hearing range, including behavior at higher frequencies (above 5,000 Hz), that traditionally has been ignored in voice research. Second, we present some environmental factors that can affect aural feedback as well as perception of voice in a typical clinical environment. Finally, we put these two discussions into context important to the voice and speech clinician.

Overlapping the Spectral Range of the Voice and the Ear

The range of the voice usually is presented in terms of the phonetogram or voice range profile (VRP). The VRP is a map of the dynamic (level) and fundamental frequency ranges of a voice. It usually is obtained using a sound level meter and a pitch source (pitch pipe or keyboard) to cue fundamental frequency, though automated techniques do exist (recommendations for obtaining VRP have been given by Schutte & Seidner, 1983). Subjects/patients match a given pitch at both a low and high intensity with various steady vowels (e.g., /i/, /a/, and /u/). A VRP is often used to show changes with treatment (Baken & Orlikoff, 2000) or vocal training (Sulter, Schutte, & Miller, 1995). Hunter, Svec, and Titze (2006) presented the VRP of a set of normal and trained vocalists; the averages from these groups are presented as examples in Figure 1.

Figure 1. Example VRP (in dB SPL) is shown for (Left): four untrained voice users, and (Right) four trained singers. Solid line is the statistical average of the people; the dotted line is the maximum and minimum across subjects. Note: From "Comparison of the Produced and Perceived Voice Range Profiles in Untrained and Trained Classical Singers," by E. J. Hunter, J. G. Svec, and I. R. Titze, 2006, Journal of Voice, *20, p. 521. Copyright 2006 by Elsevier. Reprinted with permission.*

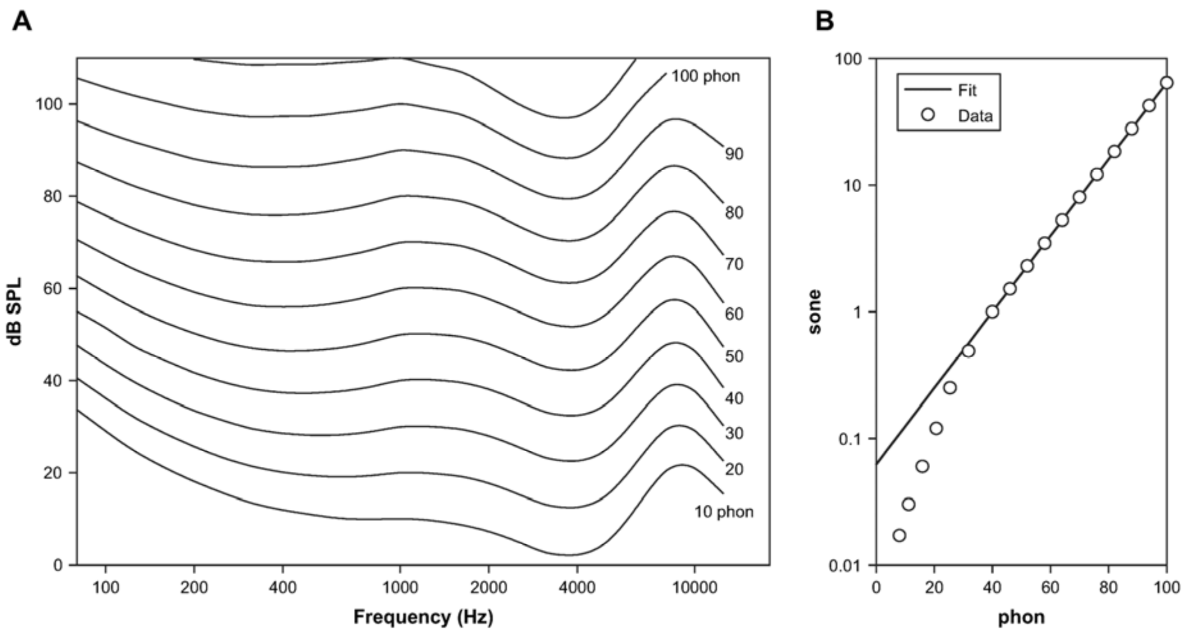


Klingholz and Fritze (1990), early researchers in using the VRP for voice classification, provided an important tutorial for measuring and interpreting the VRP of various singers. They introduced the notion of overlapping the Stimmfeld (voice field or VRP) and the Hörfeld (hearing field). However, in contrasting the voice and hearing ranges, a VRP provides only the range of the fundamental frequency of the voice, rather than the range of spectral energies produced by the voice, fundamental frequency being only a small component of the spectral range made by the voice. In order to truly compare the voice to the ear, a more complete spectrum of voice production is needed.

Before proceeding, we must briefly review the spectral sensitive nature of the ear. Human hearing sensitivity varies with frequency, as shown by the equal-loudness contour standard based on the judgment of equal loudness of pure tones (Figure 2a; ISO 226, 2003). For a given contour (an equal phon value), any frequency along that contour would be judged as having equal loudness. The dB SPL of a pure tone for a frequency can be mapped to the phon scale using the equal-loudness contours; this phon scale is equal to the dB SPL scale at 1 kHz. Thus, a 50-phon tone (regardless of the frequency) then is judged as loud as a 50 dB SPL 1 kHz tone. The phon scale illustrates that the sensitivity of the ear decreases when the frequency goes below 500 Hz and above 4 kHz (i.e., pitches lower than 500 Hz and higher than 4 KHz require a larger dB SPL for equal loudness). The Hörfeld can be thought of as the range between the threshold of hearing (approximately 0 phon) and the threshold of feeling (around 110-120 phon).

The commonly used A- and C- (and the seldom used B- and D-) weightings are attempts to simplify these equal loudness contours. However, despite the usefulness of the phon scale, it is not practical for comparing loudness values of two tones, because the phon scale is not linear (i.e., 100 phon is not twice as loud as 50 phon). Therefore, the sone scale was created (ISO 226, 2003; Kinsler, Frey, Coppens, & Sanders, 1999) to provide a linear scale for loudness (e.g., 10 sone is twice as loud as 5 sone), with the graphical relationship between phon and sone shown in Figure 2b. (Refer to ISO 226 for a more complete look at perception of sounds including complex sounds.)

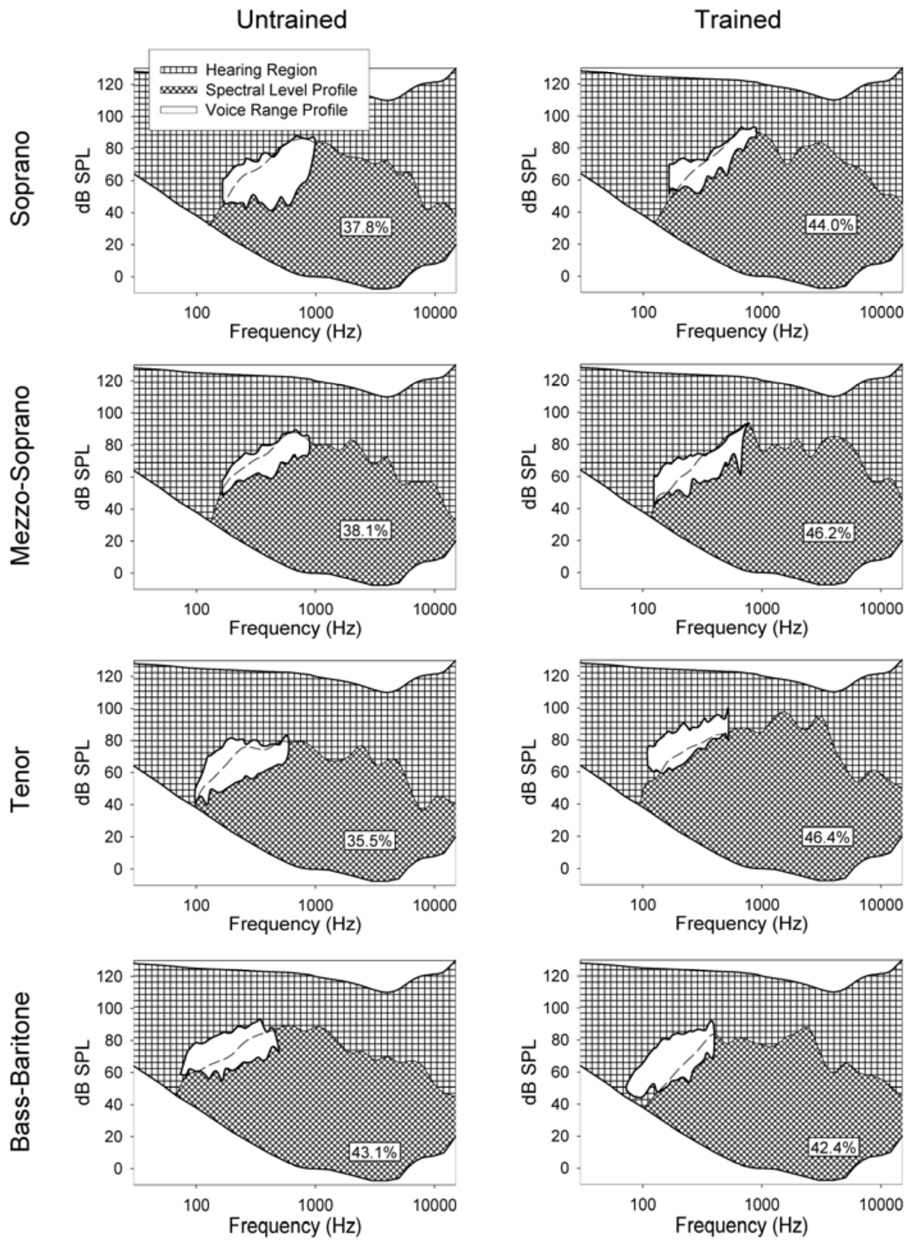
Figure 2. (a) equal loudnesss contours of the ear, with the hearing range approximately being between the 0 phon to 110 phon. (b) relation between phon and sone (a scale of loudness). Note: From "Comparison of the Produced and Perceived Voice Range Profiles in Untrained and Trained Classical Singers," by E. J. Hunter, J. G. Svec, and I. R. Titze, 2006, Journal of Voice, *20, p. 516. Copyright 2006 by Elsevier. Reprinted with permission.*



In order to compare the voice spectral range and the hearing range, Hunter and Titze (2004) presented a third-octave band spectrum analysis (in dB SPL) from the vocal output used in the VRP creation from Figure 1. These results were plotted as a Spectral Level Profile (Figure 3, diagonal cross-hatch), with the VRP (small white areas). Where the VRP overlaps the Spectral Level Profile, the edge is depicted by a dashed line. These were both superimposed on the hearing range. Percentage overlap of the hearing region and the Spectral Level Profile is shown in the figures.

In general, the untrained vocalists' Spectral Level Profiles were in closest proximity to the discomfort level of hearing (top border) at approximately 600 Hz, 30-40 dB from the threshold; this is near the first formant frequency of vowels. While the trained vocalists' Spectral Level Profiles also had a similar peak, they had a second peak nearer to the discomfort level, about 3100 Hz, with a difference of between 20-25 dB. This is the spectral region of the singer's formant, a spectral enhancement utilized by classical operatic singers where the epilarynx is used as a quarter length resonator to clustering formants 3, 4, and 5 (e.g., Yanagisawa, et al. 1989). Therefore, the singer's formant not only provides a spectral boost, but is perceptually enhanced because of its proximity to the discomfort level (the ear's most sensitive region, around 3-4 kHz). Further, trained vocalists used, on average, 45% of the hearing range (at 1 meter) compared to 38% for untrained vocalists. In general, the Spectral Level Profile is similar to the long-term average spectra of voice (Oliveira-Barrichelo, Heuer, Dean, & Sataloff, 2001; Tanner, Roy, Ash, & Buder, 2005).

Figure 3. Hearing region with overall maximum Spectral Level Profile (diagonal cross-hatch) and voice range profile (VRP) (white area) plotted on top. The percentage of the hearing range overlapped by the Spectral Level Profile was shown in a box. Note: From "Overlap of Hearing and Voicing Ranges in Singing," by E. J. Hunter and I. R. Titze, 2004, Journal of Singing, 61, p. 309. Copyright 2004 by Journal of Singing. *Reprinted with permission.*



Examination of the Spectral Level Profile (Figure 3) shows a rapid fall-off in energy above 5,000 Hz, more prominent in men's voices than in women's. Generally, the energy above 5,000 Hz (the high-frequency energy in the human voice spectrum, or HFE) has largely been ignored in voice research. While we do not understand all of the generation mechanisms of HFE, we know it can consist of (a) flow noise generated during production of consonants and/or vowels (Zhang, Mongeau, & Frankel, 2002; Zhang & Mongeau, 2006) or (b) upper harmonics of the fundamental frequency of vibration during voicing (Ternstrom, 2008). In general, the flow noise of consonants (usually fricatives) tends to dominate the overall energy level of HFE in speech. However, of particular interest to the voice community, Ternstrom found that harmonic energy from voicing can be present up to 20 kHz (the upper limit of the hearing range) in loud singing. This finding contradicts assumptions that have been held for years in the field of voice science.

Because it overlaps the human hearing region (see Figure 3), HFE affects our perception of voice. Preliminary research has suggested that listeners are highly sensitive to subtle changes in HFE, even in samples of isolated voice (i.e., with no flow noise from consonants; Monson, Lotto, & Ternstrom, 2010). Moore and Tan (2003) discovered that removing all of the energy above 7 kHz markedly decreased the "naturalness" scores of speech recordings. Perhaps of greater interest from a clinical standpoint is the effect HFE may have on the percept of "breathiness." Some researchers have proposed the use of an acoustic measure termed the "high-frequency ratio" for vocal assessment (Hartl, Hans, Vaissiere, & Brasnu, 2003; Shoji, Regenbogen, Yu, & Blaugrund, 1992). This measure is the ratio of spectral energy from 6-10 kHz to overall spectral energy from 0-10 kHz. This ratio has been found to differ significantly between populations perceived as normal and those perceived as breathy. This being the case, it is not unlikely that the human auditory nervous system would attend to this frequency range when given the task of characterizing a voice as breathy or non-breathy. It is also possible that HFE contributes to other percepts germane to clinical assessment (e.g., roughness or strain), though studies looking specifically at these relations have not been performed.

Perceiving Voice, Environmental Factors

Acoustic feedback (what we hear) is dependent on the environment where the sounds are created and perceived. The acoustic environment will affect the sound heard by a listener, and it may also cause the talker to modify her/his voice production. Thus, a review of some acoustic environmental contaminations in a clinical setting is important.

We will assume that a traditional clinic room is relatively small. Four primary acoustic (wave) effects may contaminate the sound of interest. First, sound wave reflection is the effect of sound bouncing off an object or interface where there is a change of acoustic impedance (e.g., wall, window). Second, transmission is the acoustic wave effect of sound traveling through a barrier (e.g., internal sound energy leaving the room or external sound energy coming into the room). Third, absorption is the effect of acoustic energy lost during a reflection of sound off, for example, a soft material. Finally, diffusion is the scattering of sound waves as they reflect off an uneven surface (similar to hazy or uneven glass diffusing light). These effects may be enhanced or attenuated, depending on the construction of the room. More detailed descriptions of these effects can be found in most acoustic textbooks.

An important acoustic contaminate to consider in clinical settings is reverberation, products of reflection and diffusion effects. While small rooms (like clinic rooms) generally will have lower reverberation time than larger rooms, reverberation time (RT60) should be no longer than about 0.5 seconds for good intelligibility. If it is much longer, speech sounds will be competing with preceding spoken sounds still reverberating in the room. This contaminate can be reduced by increasing absorption by introducing soft materials, like carpet or acoustical ceiling tiles. Note that these materials are most effective when distributed throughout a room, rather than concentrated on just one surface. Further, different types of materials may absorb frequencies in sound differently so even in a seemingly reverberant controlled room, an uncontrolled frequency band may still reverberate and color the resulting sound.

Flutter echoes (another product of reflection) also may interfere with intelligibility. Flutter echoes are created when a sound reflects between two parallel, hard, flat surfaces or walls, creating a flutter or ringing. Flutter echoes can be checked by simply producing a single clap in the space to see if a ringing of the clap occurs. To control them, absorption can again be used, focusing the material in a single location to minimize the parallel wall reflections. In addition to absorptive material, diffusers or objects that are highly non-flat can be used.

Finally, there is almost always some amount of noise (or unwanted sound) in the environment, which can lead to intelligibility loss or even mischaracterization of a sound. In the clinic room, the source may be from internal noises, like equipment fans from an endoscopic light box or computer. Unwanted sound also may stem from external noises like road traffic, ventilation systems, or clinic public address systems. In such cases, one should first reduce the noise at the source, if possible. If this is impossible, then isolation and distance may be used. Isolation may be improved by eliminating holes, gaps, or cracks in any interfaces between the sound source and the listening area (e.g., check window seals or introduce rubber edges at bottom and sides of doors). Further, increasing the distance from the source of noise and/or decreasing distance to the sound of interest is also valuable (in an open space, a two-fold change in distance from a source is 3dB change in level). These distance adjustments can be encapsulated in the signal-to-noise ratio, which is the difference in dB at a particular location between a sound of interest (signal) and the level of the noise. Good speech intelligibility requires about a +10dB signal-to-noise ratio (Adams & Moore, 2009; Dubbelboer & Houtgast, 2008).

Hearing Voice in a Clinical Environment

The sections above have implications that may affect the clinician's perception of a patient's voice during assessment. They also may affect a patient's self-perception of his/her voice. We will discuss these below.

Clinician's Perception of the Patient

Because of the importance of a clinician's auditory system, it behooves every clinician to have his or her hearing assessed, with particular attention paid to the high frequency range where personal sensation of hearing loss may not be noticeable. HFE plays a role in perception of certain voice qualities (e.g., "breathiness"); thus, even hearing loss at high frequencies (which is normal as a person ages) might affect a clinician's ability to accurately detail a voice disorder.

Spectral Level Profile and Hearing Range Overlap

Hunter and Titze (2004) showed that, in an anechoic environment, the Spectral Level Profile would increase its area overlap of the hearing range by about 9% when reducing the distance from a vocal performer from 1 meter to 30 cm. While this distance may not be feasible for a clinical assessment, the clinician ought to attempt to decrease the listening distance, thereby increasing the spectral overlap and presumably giving the clinician more spectral information to make an accurate assessment.

Concerning environmental factors, it has been shown that excessive noise and reverberation interfere with speech intelligibility and affect a clinician's ability to both understand and rate a patient's voice (Zraick et al., in press), especially in light of the advent of teletherapy (Xue & Lower, 2010). Common ambient noise (e.g., fan) may mask perceived naturalness or breathiness of speech during initial review and treatment (Bradley, Reich, & Norcross, 1999; Larsen & Blair, 2008; Yang & Bradley, 2009). For these reasons, clinicians should inspect the acoustical environment within which they are performing clinical assessments. For example, if the room is a particularly noisy or reverberant room, efforts ought to be made to treat it using the acoustical principles discussed in the previous section. Alternatively, decreasing the listening distance during assessment will help to improve the signal-to-noise ratio.

Patient's Self-Perception

Voice habilitation and rehabilitation revolve around the use of aural feedback as a primary source of information in an effort to train what one feels in making the acoustic output. This suspected vocal/auditory loop probably explains why several voice training/rehabilitation techniques rely on the patient's ability to hear subtle vocal changes at some level of accuracy and precision during some point in the process. For example, many voice training and/or treatment techniques focus on training a feeling based on "targeted" vocal output (e.g., resonant voice therapy, the Accent Method, Yawn-Sigh, Y-Buzz, and Vocal

Function Exercises; Barrichelo-Lindstrom & Behlau, 2009; Dalhoff & Kitzing, 1977; Griffin, Woo, Colton, Casper, & Breuder, 1995; Kotby, Shiromoto, & Hirano, 1993; Sabol, Lee, & Stemple, 1995; Stemple, Lee, D'amico, & Pickup, 1994; Stemple, 2005; Titze, 2001).

People use self-perceived vocal output to help them make real time changes to their voices to control pitch and loudness for prosody and other suprasegmentals for communication (e.g., speech rate, pitch variation and perturbation, articulation precision, word stress/emphasis, and voice quality; Bauer, Mittal, Larson, & Hair, 2006; Fairbanks & Guttman, 1958; Xue & Lower, 2010). Environment acoustic factors also can affect vocal output. For example, people often raise their voices if the background noise is raised (Garnier, Henrich, & Dubois, 2010; Patel & Schell, 2008). This loudness change also may raise the phonation pitch (Gramming, Sundberg, Ternstrom, Leanderson, & Perkins, 1988).

The person with whom one is communicating also can affect one's vocal output. For example, it has been shown that children perform voice differently depending on vocal task type (Baker, Weinrich, Bevington, Scroth, & Schroeder, 2008) and communication environment (Hunter, 2009). Human speech/voice perception is such that one learns to distinguish variations in another person's voice to make judgments regarding their emotional status (e.g., happy, sad, bored, interested) in as quickly as 200 ms (Paulmann & Kotz, 2008). All of these perceptions then can change one's own vocal patterns, as part of a natural reaction to his/her own emotional status.

Thus, talkers use their own acoustic feedback as a test for their intended vocal success. Because of the importance of this acoustic feedback, the acoustic environment of a patient (including the clinician's own communication to the patient) ought to be considered and modified, if necessary, so that the patient's voice production is not habitually produced to allow for accurate assessment of the patient's voice.

Conclusion

The ear and the vocal mechanism are closely tied. Because of this relationship, several factors that affect the hearing of a listener also can affect the listener's perception of the talker. Likewise, the hearing of the talker can affect his or her own vocal output. It is valuable for a clinician to understand these principles when attempting to accurately assess the voice of a patient in the clinic. In the event that the hearing of either the clinician or patient is degraded in some fashion (e.g., hearing loss, environmental acoustical interference), measures can and should be taken to remedy this situation wherever possible.

Acknowledgment

This work was supported by grants R01-DC04347 and F31-DC010533 from the National Institutes of Health/National Institute on Deafness and Other Communication Disorders.

References

Adams, E. M., & Moore, R. E. (2009). Effects of speech rate, background noise, and simulated hearing loss on speech rate judgment and speech intelligibility in young listeners. *Journal of the American Academy of Audiology, 20,* 28-39.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular Publishing Group.

Baker, S., Weinrich, B., Bevington, M., Schroth, K., & Schroeder, E. (2008). The effect of task type on fundamental frequency in children. *International Journal of Pediatric Otorhinolarngology, 72,* 885-889.

Barrichelo-Lindstrom, V., & Behlau, M. (2009). Resonant voice in acting students: Perceptual and acoustic correlates of the trained Y-Buzz by Lessac. *Journal of Voice, 23,* 603-609.

Bauer, J. J., Mittal, J., Larson, C. R., & Hain, T. C. (2006). Vocal responses to unanticipated perturbations in voice loudness feedback: An automatic mechanism for stabilizing voice amplitude. *Journal of the Acoustic Society of America, 119,* 2363-2371.

Bradley, J. S., Reich, R. D., & Norcross, S. G. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. *Journal of the Acoustic Society of America, 106,* 1820-1828.

Dalhoff, K. & Kitzing, P. (1977). [Observations on the accent-method of Smith in the management of speech and voice disturbances. Part I (author's transl)]. *HNO, 25*, 102-105.

Dubbelboer, F., & Houtgast, T. (2008). The concept of signal-to-noise ratio in the modulation domain and speech intelligibility. *Journal of the Acoustic Society of America, 124,* 3937-3946.

Fairbanks, G., & Guttman, N. (1958). Effects of delayed auditory feedback upon articulation. *Journal of Speech and Hearing Research, 1*, 12-22.

Garnier, M., Henrich, N., & Dubois, D. (2010). Influence of sound immersion and communicative interaction on the Lombard effect. *Journal of Speech and Hearing Research, 53,* 588-608.

Gramming, P., Sundberg, J., Ternstrom, S., Leanderson, R., & Perkins, W. (1988). Relationship between changes in voice pitch and loudness. *Journal of Voice, 2,* 118-126.

Griffin, B., Woo, P., Colton, R., Casper, J., & Brewer, D. (1995). Physiological characteristics of the supported singing voice: A preliminary study. *Journal of Voice, 9*, 45-56.

Hartl, D. A., Hans, S., Vaissiere, J., & Brasnu, D. A. (2003). Objective acoustic and aerodynamic measures of breathiness in paralytic dysphonia. *European Archives of Otorhinolaryngology, 260,* 175-182.

Hunter, E. J. (2009). A comparison of a child's fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study. *International Journal of Pediatric Otorhinolaryngology, 73*, 561-571.

Hunter, E. J., Svec, J. G., & Titze, I. R. (2006) Comparison of the produced and perceived voice range profiles in untrained and trained classical singers. *Journal of Voice, 20,* 513-526.

Hunter, E. J., & Titze, I. R. (2004). Overlap of hearing and voicing ranges in singing. *Journal of Singing, 61*, 387-392.

International Organization for Standardization. (2003). *Acoustics—Normal equal-loudness-level contours.* Geneva, Switzerland: Author.

Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (1999). *Fundamentals of acoustics* (4th ed.). Hoboken, NJ: Wiley Text Books.

Klingholz, F. (1990). *The voice field: A practical guide for measurement and evaluation* (in German). Munich, Germany: Verlag J. Peperny.

Kotby, M. N., Shiromoto, O., & Hirano, M. (1993). The accent method of voice therapy: Effect of accentuations on FO, SPL, and airflow. *Journal of Voice, 7,* 319-325.

Larsen, J. B., & Blair, J. C. (2008). The effect of classroom amplification on the signal-to-noise ratio in classrooms while class is in session. *Language Speech and Hearing Services in the Schools, 39*, 451-460.

Monson, B. M., Lotto, A. J., & Ternstrom, S. (2010). *Detection of high-frequency energy changes in sustained vowels produced by singers.* Manuscript submitted for publication.

Moore, B. C., & Tan, C. T. (2003). Perceived naturalness of spectrally distorted speech and music. *Journal of the Acoustic Society of America, 114,* 408-419.

Oliveira-Barrichelo, V., Heuer, R. J., Dean, C. M., & Sataloff, R. T. (2001). Comparison of singer's formant, speaker's ring, and LTA spectrum among classical singers and untrained normal speakers. *Journal of Voice, 15,* 344-350.

Patel, R., & Schell, K. W. (2008). The influence of linguistic content on the Lombard effect. *Journal of Speech, Language, and Hearing Research, 51*, 209-220.

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *Neuroreport, 19*, 209-213.

Sabol, J. W., Lee, L., & Stemple, J. C. (1995). The value of vocal function exercises in the practice regimen of singers. *Journal of Voice, 9,* 27-36.

Schutte, H. K., & Seidner, W. (1983). Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement/phonetography. *Folia Phoniatrica (Basel), 35,* 286-288.

Shoji, K., Regenbogen, E., Yu, J. D., & Blaugrund, S. M. (1992). High-frequency power ratio of breathy voice. *Laryngoscope, 102*, 267-271.

Stemple, J. C. (2005). A holistic approach to voice therapy. *Seminars in Speech and Language, 26*, 131-137.

Stemple, J. C., Lee, L., D'Amico, B., & Pickup, B. (1994). Efficacy of vocal function exercises as a method of improving voice production. *Journal of Voice, 8,* 271-278.

Sulter, A. M., Schutte, H. K., & Miller, D. G. (1995). Differences in phonetogram features between male and female subjects with and without vocal training. *Journal of Voice, 9*, 363-377.

Tanner, K., Roy, N., Ash, A., & Buder, E. H. (2005). Spectral moments of the long-term average spectrum: sensitive indices of voice change after therapy? *Journal of Voice, 19*, 211-222.

Ternstrom, S. (2008). High fidelity on the line: please say "aah." *Echoes Newsletter of the Acoustic Society of America 18,* 4-5.

Titze, I. R. (2001). Acoustic interpretation of resonant voice. *Journal of Voice, 15,* 519-528.

Xue, S. A., & Lower, A. (2010). Acoustic fidelity of internet bandwidths for measures used in speech and voice disorders. *Journal of the Acoustic Society of America, 128,* 1366-1376.

Yang, W., & Bradley, J. S. (2009). Effects of room acoustics on the intelligibility of speech in classrooms for young children. *Journal of the Acoustic Society of America, 125*, 922-933.

Yanagisawa, E. Estill, J., Kmucha, T., & Leder, S. B. (1989). The contribution of aryepiglottic constriction to "ringing" voice quality-A videolaryngoscopic study with acoustic analysis. *Journal of Voice, 3*, 342-350.

Zhang, Z., Mongeau, L., & Frankel, S. H. (2002). Broadband sound generation by confined turbulent jets. *Journal of the Acoustic Society of America, 112*, 677-689.

Zhang, Z., & Mongeau, L. G. (2006). Broadband sound generation by confined pulsating jets in a mechanical model of the human larynx. *Journal of the Acoustic Society of America, 119*, 3995-4005.

Zraick, R. I., Kempster, G. B., Connor, N. P., Thibeault, S., Klaben, B. K., Bursac, Z. et al. (in press). Establishing validity of the Consensus Auditory Perceptual Evaluation of Voice (CAPE-V). *American Journal of Speech-Language Pathology.*